

# DOI Background Information

## Introduction to DOIs

A Digital Object Identifier or DOI is a unique alphanumeric string used to identify a digital object and provide a permanent link online. DOIs are often used in online publications in citations. DOIs are assigned and regulated by The International DOI Foundation (IDF). ()

DOIs are alphanumeric strings in the following format:

–doi:[prefix]/[suffix]

Prefix – 10.[number] where [number] identifies registrant agent. 5067 has been assigned for NASA ESDIS.

Suffix – uniquely identifies the data item and its format is assigned and managed by the registrant agent.

DOIs provide persistent identification, enabling easier access to research data. They have additional value when stored in product metadata. Researchers that acquire product data files should be able use the DOI to find the definitive documentation from NASA's Scientific and Technical Information archives. Adding DOIs to product metadata will also enable tools for provenance tracking – help find more information about the creation of the data product. This increases verification and validation of scientific results. DOIs can increase the potential for scientific data citations.

DOIs are resolved by entering the full DOI name into <http://dx.doi.org/>. Entering the DOI name in the "doi" and clicking "Go" will take the user to the product landing page that corresponds to the product that goes with that DOI.

- DOI Background Information
- ESDIS DOI Process
- DOI Submission Process
- DOI Landing Page
- Contact Information
- EOSDIS DOIs Status and Listing
- DOI Documents
- ESDSWG Recommendations
- FAQs
- References and Links

## DOI Citation

There is an online tool that can generate a citation by entering the doi and selecting the appropriate journal and language.

<http://crosscite.org/citeproc/>

For example, the 10.5067/AURA/HIRDLS/DATA101 for the American Geological Union in American English:

Zwally, H. J. and Schutz, Bob (2012), GLAS/ICESat L1B Global Elevation Data (HDF5), , doi:10.5067/ICESAT/GLAS/DATA106. [online] Available from: <http://dx.doi.org/10.5067/ICESAT/GLAS/DATA106>

## Putting DOIs in Product Metadata

DOIs have additional value when stored in product metadata. Researchers that acquire product data files should be able use the DOI to find the definitive documentation from NASA's Scientific and Technical Information archives. Adding DOIs to product metadata will also enable tools for provenance tracking – help find more information about the creation of the data product.

Existing metadata structures can accommodate the addition of EOSDIS product and file level identifiers in multiple ways. The purpose of this technical review is to examine various aspects of each and a combination of multiple implementations.

The product data files will contain a file level attribute for the DOI. The global file-level attribute is required for the user convenience (e.g., viewing with standard exploratory tools).

The associated EOSDIS core metadata and metadata files should also contain a Product Specific Attribute (PSA) for the DOI (e.g., to enable applications).

## DOI Attribute Name

The tag to use in the metadata, i.e., the attribute name, will be "identifier\_product\_doi". This syntax was chosen to support a standard way of handling multiple kinds of identifiers, and both file-level and product-level identifiers. The same attribute name will be used for both file-level (i.e., global) and PSA constructs.

We are able to set DOI values in both the collection-level and granule-level metadata for the same

ESDT (Earth Science Data Type, structure used for ingesting products into ECS archives). A value for the "identifier\_product\_doi" PSA will be coded into the collection-level metadata section of an ESDT descriptor file so that it will be exported to ECHO. Adding a value for the "identifier\_product\_doi" PSA is a simple collection-level metadata change in the ESDT descriptor file, and can be done after the ESDT's granules have been produced.

The production software will put the "identifier\_product\_doi" PSA and its value in the granule metadata files and in the HDF files.

Data Product Attribute	Definition	Allowed Values
Identifier	DOI Name, the Identifier which is a unique string that identifies a resource.	DOI name
Creator Name	The name of person(s) or data center who created the product.	List of name(s)
Title	Product name, a name or title by which a product is known.	Title of the product
Publisher	The name of the data provider that archives and/or distributes the data products.	Name of the organization that distributes the data.
Publication Year	The year when the data was or will be made publicly available.	YYYY, four digits of the year
URL	URL for the landing page that has the current location of the information of the data product.	Web URL format

## DOI Attribute Name Authority

An additional attribute would define the authoritative service for use with DOI values in resolving to the URL location. Adding this attribute to HDF metadata will allow more complete mapping to netCDF and to ISO standard metadata structures.

The most valuable place for this attribute is in the HDF files, so it should be added at the file level as well as the PSA level.

It is best if the attribute name for the DOI authority is similar to the name for the DOI identifier i.e., similar to "identifier\_product\_doi" so we are asking providers to add an attribute name = "identifier\_product\_doi\_authority". This would help users (and code) find it in the metadata.

The value of the attribute would be "<http://dx.doi.org/>"

## Lower case attribute naming convention for netCDF CF

It has been pointed out that the common practice for netCDF CF is to define attribute names in all lower case, using underscores to separate key words. This has been generally extended for non-standard attributes (such as 'identifier\_product\_doi'). If upper case letters are used there will need to be special code in translators and testing with netCDF tools. So as of this writing, we want to use an all lower case attribute name "identifier\_product\_doi" and attribute naming authority "identifier\_product\_doi\_authority".